Commutator maps, measure preservation, and T-systems

Shelly Garion and Aner Shalev
Institute of Mathematics
Hebrew University
Jerusalem 91904
Israel

February 1, 2008

Abstract

Let G be a finite simple group. We show that the commutator map $\alpha : G \times G \to G$ is almost equidistributed as $|G| \to \infty$. This somewhat surprising result has many applications. It shows that for a subset $X \subseteq G$ we have $\alpha^{-1}(X)/|G|^2 = |X|/|G| + o(1)$, namely α is almost measure preserving. From this we deduce that almost all elements $g \in G$ can be expressed as commutators $g = [x, y]$ where x, y generate G.

This enables us to solve some open problems regarding T-systems and the Product Replacement Algorithm (PRA) graph. We show that the number of T-systems in G with two generators tends to infinity as $|G| \to \infty$. This settles a conjecture of Guralnick and Pak. A similar result follows for the number of connected components of the PRA graph of G with two generators.

Some of our results apply for more general finite groups, and more general word maps.

Our methods are based on representation theory, combining classical character theory with recent results on character degrees and values in finite simple groups. In particular the so called Witten zeta function $\zeta^G(s) = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^{-s}$ plays a key role in the proofs.

This article was submitted to the *Transactions of the American Mathematical Society* on 21 February 2007 and accepted on 24 June 2007

The second author acknowledges the support of grants from the Israel Science Foundation and the Bi-National Science Foundation United-States Israel

2000 *Mathematics Subject Classification:* 20D06, 20P05, 20D60

1 Introduction

1.1 Finite groups

Let G be a finite group. Let $\alpha = \alpha_G : G \times G \to G$ be the commutator map, namely

$$\alpha(x, y) = [x, y] = x^{-1}y^{-1}xy.$$

How equidistributed is this map?

To make the question more precise, define for $g \in G$

$$N(g) = |\alpha^{-1}(g)|,$$

the size of the fiber above g. When can we show that $N(g)$ is roughly $|G|$ for almost all $g \in G$?

For general groups this is often far from true. However, we shall show below that commutator maps on finite simple groups are almost equidistributed. More generally, we associate with each finite group G a certain parameter $\epsilon(G)$ related to its representation degrees, and prove that if $\epsilon(G)$ is small then α_G is almost equidistributed.

We need some notation. For a finite group G let $\mathrm{Irr}(G)$ denote the set of complex irreducible characters of G. The numbers $N(g)$ above can be studied using a character-theoretic approach, based on Frobenius classical formula

$$N(g) = |G| \sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(g)}{\chi(1)}. \tag{1}$$

Thus the character table of G provides complete information on the distribution of the commutator map. However, computing or estimating the right hand side of (1) for infinite families of groups, without complete information on their character tables, is often a formidable task. We shall see below how information on character degrees (but not character values) sometimes suffices to draw strong conclusions regarding commutator maps.

The character degrees of G are conveniently encoded in the so called *Witten zeta function* ζ^G of G, defined by

$$\zeta^G(s) = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^{-s},$$

where s is a real number. This function, which plays an important role in this paper, was originally defined and studied by Witten [Wit] for Lie groups. For finite simple groups it was studied and applied in detail in [LiSh3, LiSh4, LiSh5].

Let $P = P^G$ be the commutator distribution on G, namely

$$P(g) = N(g)/|G|^2,$$

and let U be the uniform distribution on G (so $U(g) = 1/|G|$ for all $g \in G$).

Our first result bounds the L_1-distance

$$||P - U||_1 = \sum_{g \in G} |P(g) - U(g)|$$

between the probability measures above.

Proposition 1.1 *Let G be a finite group. Then we have*

$$||P^G - U^G||_1 \leq (\zeta^G(2) - 1)^{1/2}.$$

Set

$$\delta(G) = (\zeta^G(2) - 1)^{1/2}.$$

We now deduce a general lower bound on the number of commutators in G.

Corollary 1.2 *A finite group G has at least $(1 - \delta(G))|G|$ commutators.*

Next we define

$$\epsilon(G) = (\zeta^G(2) - 1)^{1/4}.$$

We establish equidistribution properties of the commutator map on groups in terms of the parameter $\epsilon(G)$ defined above.

Theorem 1.3 *Every finite group G has a subset $S = S_G \subseteq G$ with the following properties:*

(i) $|S| \geq (1 - \epsilon(G))|G|$;

(ii) $(1 - \epsilon(G))|G| \leq N(g) \leq (1 + \epsilon(G))|G|$ for all $g \in S$.

Of course results 1.1-1.3 above have no content when $\zeta^G(2) \geq 2$ (since then $\delta(G), \epsilon(G) \geq 1$). Since the linear characters of G contribute $|G : G'|$ to $\zeta^G(2)$ we see that these results can only be useful for perfect groups, namely groups for which $G = G'$.

Recall that the *representation growth* of G is the series $\{r_n(G)\}$, where $r_n(G)$ is the number of irreducible representations of G of degree n. See [LM, Ja] for background. We clearly have

$$\zeta^G(s) = \sum_{n \geq 1} r_n(G) n^{-s}.$$

If G is perfect, and has very small representation growth, so that

$$\sum_{n \geq 2} r_n(G) n^{-2} < 1,$$

then results 1.1-1.3 can be meaningfully applied. Roughly speaking, they show that perfect groups with few representations have many commutators, and that commutator maps on them are almost equidistributed.

This in turn has some further applications.

Corollary 1.4 *Let G be a finite group.*

(i) The commutator map $\alpha = \alpha_G$ satisfies

$$\left| \frac{|\alpha^{-1}(Y)|}{|G|^2} - \frac{|Y|}{|G|} \right| \leq 3\epsilon(G) \text{ for all } Y \subseteq G.$$

(ii) If $X \subseteq G \times G$ then

$$\frac{|\alpha(X)|}{|G|} \geq \frac{|X|}{|G|^2} - 3\epsilon(G).$$

Part (i) above shows that, if $\epsilon(G)$ is close to zero, then the commutator map on G is almost measure preserving.

1.2 Finite simple groups

The main context in which the above results can be successfully applied is that of finite simple groups. Indeed, by Theorem 1.1 of [LiSh4], if G is simple, then

$$\zeta^G(s) \to 1 \text{ as } |G| \to \infty \text{ provided } s > 1.$$

It follows that for finite simple groups G, $\delta(G)$ and $\epsilon(G)$ tend to zero as $|G| \to \infty$.

Applying this we deduce our main equidistribution results for commutator maps. Theorem 1.3 gives rise to the following.

Theorem 1.5 *Every finite simple group G has a subset $S = S_G \subseteq G$ with the following properties:*

(i) $|S| = |G|(1 - o(1))$;

(ii) $N(g) = |G|(1 + o(1))$ uniformly for all $g \in S$.

Here and throughout this paper $o(1)$ denotes a real number depending only on G which tends to zero as $|G| \to \infty$.

The proofs of Theorem 1.5, and of our next results below, rely on the Classification of Finite Simple Groups. It clearly suffices to consider alternating groups A_n and simple groups of Lie type. We show in Section 4 below that $\epsilon(A_n) = O(n^{-1/2})$, and if G is of Lie type of rank r over a field with q elements then $\epsilon(G) = O(q^{-r/4})$. This provides explicit upper bounds on the error term $o(1)$ in Theorem 1.5.

Note that we cannot require in the theorem above that $S = G$. Indeed, it is well known (and follows from (1) above) that

$$N(1) = |G|k(G),$$

where $k(G)$ is the number of conjugacy classes in G. Since $k(G) \to \infty$ as $|G| \to \infty$ we see that the fiber above $g = 1$ is large and does not satisfy condition (ii).

Theorem 1.5 amounts to saying that, for a finite simple group G,

$$||P^G - U^G||_1 \to 0 \text{ as } |G| \to \infty.$$

We present two proofs of Theorem 1.5. The first is probabilistic, based on Proposition 1.1 bounding the L_1-distance above, which proves the existence of the required subsets S. In the second proof we construct the subsets S explicitly, which often yields better lower bounds on their cardinality. For example, in the constructive proof for A_n we obtain $|S| \geq |A_n|(1-2/[\sqrt{n}]!)$, which is much better than the $|A_n|(1-O(n^{-1/2}))$ lower bound given by the probabilistic proof. See Section 5 for more details.

Our constructive proof of Theorem 1.5 applies some powerful recent results on character values and degrees (see [LiSh3, LiSh4, MS, Sh]) to estimate the right hand side of Frobenius formula (1) for specific (almost all) elements $g \in G$, showing that the main contribution comes from the trivial character $\chi = 1$, and the accumulative contribution of all non-trivial characters is marginal.

Theorem 1.5 can be seen as a culmination of various results on commutators in finite simple groups. This topic has a long history. A conjecture of Ore [O] from 1951, which is still not fully resolved (see [EG]), states that all elements of a finite simple group are commutators. Various results were obtained in order to present group elements as commutators or short products of commutators, see for instance [Ga, Wi, Go, Sh].

Theorem 1.5 can be viewed as a probabilistic version of Ore's conjecture. It obviously implies that almost all elements of a finite simple group are commutators, a result obtained recently in [Sh], Theorem 2.9(i).

However, Theorem 1.5 has more refined consequences, as follows.

Corollary 1.6 *Let G be a finite simple group.*

(i) The commutator map α is almost measure preserving, namely

$$\frac{|\alpha^{-1}(Y)|}{|G|^2} = \frac{|Y|}{|G|} + o(1) \text{ for all } Y \subseteq G.$$

(ii) If $X \subseteq G \times G$ then

$$\frac{|\alpha(X)|}{|G|} \geq \frac{|X|}{|G|^2} - o(1).$$

(iii) In particular, if $X \subseteq G \times G$ satisfies $|X| = (1-o(1))|G|^2$ then

$$|\alpha(X)| = (1-o(1))|G|,$$

namely almost all elements of G can be represented as commutators $[x,y]$ where $(x,y) \in X$.

Indeed, this follows by combining Corollary 1.4 with the fact that $\epsilon(G) = o(1)$.

As a consequence of (ii) above we see that if A, B are subsets of G such that $|A| \geq a|G|$ and $|B| \geq b|G|$, then the set $[A,B]$ of commutators $[x,y]$ where $x \in A, y \in B$ satisfies

$$|[A,B]| \geq (ab - 3\epsilon(G))|G| = (ab - o(1))|G|.$$

An old conjecture of Dixon [Di], which is now a theorem (see [B], [KL], [LiSh1]), states that almost all pairs of elements of a finite simple group are generating pairs. Applying Corollary 1.6 for the set X of generating pairs of G we obtain

Theorem 1.7 *Let G be a finite simple group, and let $g \in G$ be randomly chosen. Then the probability that g can be represented as a commutator $g = [x,y]$ where x, y generate G tends to 1 as $|G| \to \infty$.*

While Ore's conjecture has been established for alternating groups and for simple groups of Lie type over large fields, Theorem 1.7 is new for all types of finite simple groups.

Combining a character-theoretic approach with probabilistic arguments enables us to obtain additional equidistribution results (see Theorems 7.1 and 7.4 below). We show that maps on finite simple groups induced by the word x^2y^2, or by longer commutators in any arrangement of brackets, are almost equidistributed.

1.3 T-systems and the Product Replacement Algorithm

A main motivation behind Theorem 1.7 comes from the study of transitivity systems, also known as T-systems.

Let G be a finite group and let $d(G)$ be the minimal number of generators of G. For $k \geq 1$ let F_k denote the free group on k generators. For any $k \geq d(G)$, let

$$V_k(G) = \{(g_1, \ldots, g_k) \in G^k : \langle g_1, \ldots, g_k \rangle = G\}$$

be the set of all generating k-tuples of G. One can identify $V_k(G)$ with the set of epimorphisms $\mathrm{Epi}(F_k \to G)$. The group $\mathrm{Aut}(F_k) \times \mathrm{Aut}(G)$ acts on $V_k(G)$ by $(\tau, \sigma) : \phi \to \sigma \circ \phi \circ \tau^{-1}$, where $\tau \in \mathrm{Aut}(F_k), \sigma \in \mathrm{Aut}(G)$ and $\phi \in \mathrm{Epi}(F_k \to G)$. The orbits of this action are called *systems of transitivity*, and also *T-systems* or *T_k-systems*, when we specify the value of k. They were introduced

by B.H. Neumann and H. Neuman in [NN] in the context of presentations of finite groups and studied further in [Du1, Du2, E2, Gi, GP, Ne, P].

It is well-known that $\mathrm{Aut}(F_k)$ is generated by the following moves, called the *Nielsen moves*, on the k standard generators $x_1, \ldots, x_k$ of F_k, viewed as elements in $\mathrm{Aut}(F_k)$.

$$\begin{aligned}
R_{i,j} &: (x_1, \ldots, x_i, \ldots, x_k) \to (x_1, \ldots, x_i \cdot x_j, \ldots, x_k), \\
L_{i,j} &: (x_1, \ldots, x_i, \ldots, x_k) \to (x_1, \ldots, x_j \cdot x_i, \ldots, x_k), \\
P_{i,j} &: (x_1, \ldots, x_i, \ldots, x_j, \ldots, x_k) \to (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k), \\
I_i &: (x_1, \ldots, x_i, \ldots, x_k) \to (x_1, \ldots, x_i^{-1}, \ldots, x_k), \\
&\text{for } 1 \leq i \neq j \leq k.
\end{aligned}$$

These moves define the following graph: its vertices are $V_k(G)$ and its edges correspond to the Nielsen moves and to the automorphisms of G. The connected components of this graph are exactly the T_k-systems.

In recent years there is renewed interest in T-systems due to exciting applications to the Product Replacement Algorithm.

The *Product Replacement Algorithm (PRA)* is a practical algorithm to construct random elements of a finite group. The algorithm was introduced and analyzed in [CLMNO], where the authors proved that it produces asymptotically uniformly distributed elements. As the success of the algorithm became widely acknowledged, it was included as a standard routine in the two major algebra packages GAP and MAGMA. Since then the algorithm was widely investigated (see [BP, GaP, LP, P]).

The product replacement algorithm is defined as follows [CLMNO, P]. Given a generating k-tuple $(g_1, \ldots, g_k) \in V_k(G)$, a *move* to another such tuple is defined by first drawing uniformly a pair (i, j) with $1 \leq i \neq j \leq k$ and then applying one of the following four operations with equal probability:

$$\begin{aligned}
R_{i,j}^{\pm} &: (g_1, \ldots, g_i, \ldots, g_k) \to (g_1, \ldots, g_i \cdot g_j^{\pm 1}, \ldots, g_k) \\
L_{i,j}^{\pm} &: (g_1, \ldots, g_i, \ldots, g_k) \to (g_1, \ldots, g_j^{\pm 1} \cdot g_i, \ldots, g_k).
\end{aligned}$$

To produce a random element in G, start with some generating k-tuple, apply the above moves several times, and finally return a random element of the generating k-tuple that was reached.

The moves in the PRA can be conveniently encoded by the *PRA graph* $\Gamma_k(G)$ whose vertices are the tuples $V_k(G)$, with edges corresponding to the moves $R_{i,j}^{\pm}, L_{i,j}^{\pm}$. The PRA corresponds to a random walk on this graph. However, it is usually more convenient to look at the following extended graph. The *extended PRA graph* $\tilde{\Gamma}_k(G)$ is a graph on $V_k(G)$ corresponding to the *Nielsen moves*, $R_{i,j}^{\pm}, L_{i,j}^{\pm}$ and $P_{i,j}, I_i$, for $1 \leq i \neq j \leq k$.

It is clear from the definitions that the number of T_k-systems is less or equal to the number of connected components in $\tilde{\Gamma}_k(G)$, denoted by $\tilde{\chi}_k(G)$.

In addition, if $\chi_k(G)$ denotes the number of connected components in $\Gamma_k(G)$, then $\tilde{\chi}_k(G) \leq \chi_k(G) \leq 2\tilde{\chi}_k(G)$. Moreover, if $k \geq d(G)+1$ then $\Gamma_k(G)$ is connected if and only if $\tilde{\Gamma}_k(G)$ is connected, and if $k \geq 2d(G)$, then both $\Gamma_k(G)$ and $\tilde{\Gamma}_k(G)$ are connected if and only if G has only one T_k-system (see [P]).

An interesting question is to estimate the number of T_k-systems of G as a function of k and G. It is well known that a k-generated abelian group has only one T_k-system. A first example of a nilpotent group G with more than one T_k-system, where $k = d(G)$, was given in [Ne], as an answer to a question of Gaschütz. Later, Dunwoody [Du1] proved that the number of T_k-systems of certain groups is in fact not bounded, i.e. for every k, N and p, one can find a p-group G with $d(G) = k$ such that the number of T_k-systems of G is at least N.

Particular attention was given to T-systems in finite simple groups G. Here $d(G) = 2$ and a conjecture attributed to Weigold states that for $k \geq 3$ the number of T_k-systems of G is 1. This conjecture was proven for very few families of simple groups (see Gilman [Gi] and Evans [E]). However, the case $k = 2$ seems to be different. It has been shown that the number of T_2-systems in $G = L_2(p) = PSL_2(p)$ tends to infinity as $|G| \to \infty$ (see Evans [E] and Guralnick and Pak [GP]). A similar result for $G = A_n$ was proved by Pak (see [P]). In [GP] Guralnick and Pak suggest that this might be true for all finite simple groups, and remark that different methods will have to be established in order to confirm this.

In this paper we confirm this conjecture. Moreover, we obtain explicit lower bounds on the number of T_2-systems for all finite simple groups.

For a (possibly twisted) Lie type L, not ${}^2B_2$, ${}^2G_2$ or ${}^2F_4$, define the rank $r = r(L)$ to be the untwisted Lie rank of L (that is, the rank of the ambient simple algebraic group); and for L of type ${}^2B_2$, ${}^2G_2$ or ${}^2F_4$, define $r(L) = 1, 1, 2$ respectively. Let $G_r(q)$ denote a finite simple group of Lie type of rank r over a field with q elements (in the unitary case the natural module for $G_r(q)$ is over the field with q^2 elements).

Theorem 1.8 *Let G be a finite simple group. Then the number of T_2-systems in G tends to infinity as $|G| \to \infty$.*

Moreover, this number is at least $aq^r r^{-1}(\log q)^{-2}$ when $G = G_r(q)$, and at least $n^{(1/2-\epsilon)\log n}$ when $G = A_n$.

Here a is an absolute positive constant, and $\epsilon > 0$ is arbitrary provided n is large enough (namely $n \geq f(\epsilon)$).

The second, quantitative, assertion of Theorem 1.8, is new even for $L_2(p)$ and A_n, and answers a question from [GP]. The lower bound for the number of T_2-systems in A_n which follows from the argument in [P] is about

$n/(2\log n)$. No explicit lower bound was known for $L_2(p)$. In fact the detailed lower bound for $G_r(q)$ which stems from our proof of Theorem 1.8 is somewhat better than the one stated, and has the form $(1/8 - o(1))p$ when $G = L_2(p)$.

Theorem 1.8 immediately provides information on the Product Replacement Algorithm graph in case of two generators.

Corollary 1.9 *Let G be a finite simple group. Then the number of connected components of the PRA graph $\Gamma_2(G)$ tends to infinity as $|G| \to \infty$.*

Of course the number of components above is bounded below by the number of T_2-systems, hence the lower bounds of Theorem 1.8 apply. In particular $\Gamma_2(A_n)$ has at least $n^{(1/2-\epsilon)\log n}$ connected components. For groups of Lie type our method yields a somewhat better bound, showing that $\Gamma_2(G_r(q))$ has at least $aq^r(\log q)^{-1}$ connected components (see Section 6).

The main tool in our proof of Theorem 1.8 is Theorem 1.7, which is in turn a by-product of our equidistribution theorem (1.5).

1.4 Notation and layout

Our notation is rather standard. For a finite group G let $\mathrm{Irr}(G)$ denote the set of irreducible complex characters of G. Let $k(G)$ denote the number of conjugacy classes in G. For $g \in G$ we let g^G be the conjugacy class of g in G, and let $|g|$ be the order of g. We denote by $G_r(q)$ a finite simple group of rank r over a field with q elements. An element $g \in G_r(q)$ is called regular semisimple if it is semisimple and its centralizer in the ambient algebraic group has minimal dimension r. To a function $f : X \to Y$ between finite sets we associate a probability distribution P_f on Y given by

$$P_f(y) = \frac{|f^{-1}(y)|}{|X|} \text{ where } y \in Y.$$

By a word $w = w(x_1, \ldots, x_m)$ we mean an element of the free group F_m on $x_1, \ldots, x_m$. Given a group G the word w defines a function $G^m \to G$, obtained by substitution, which we denote by α_w. The associated probability distribution P_{α_w} on G will be also denoted by $P_w(G)$. Additional notation will be introduced when needed.

Some words on the layout of this paper. In Section 2 we use Fourier techniques to prove Proposition 1.1. In Section 3 we prove results 1.2-1.4. Section 4 presents recent delicate results on simple groups which are required in order to prove Theorem 1.5 constructively. The constructive proof of this theorem is then carried out in Section 5. Section 6 is devoted to the various applications, focusing on T-systems; this is where results 1.7-1.9 are

proved. In Section 7 we obtain additional equidistribution results for longer commutators.

Acknowledgments. We are grateful to Benjy Weiss for useful remarks. This paper is part of the first author's Ph.D. thesis done under the supervision of Alex Lubotzky.

2 Bounding the L_1-distance

The main purpose of this section is to prove Proposition 1.1. We shall start with a more general discussion, from which this result will be deduced.

Let G be a finite group, and let P be a probability distribution on G which is a class function. For example, we may have $P = P_w$ where w is a word.

Consider the non-commutative Fourier expansion

$$P = |G|^{-1} \sum_{\chi \in \mathrm{Irr}(G)} a_\chi \chi,$$

with suitable (complex) coefficients a_χ.

Note that $a_1 = 1$ (since $\sum_{g \in G} P(g) = 1$).

Lemma 2.1 *We have*

$$\sum_{g \in G} P(g)^2 = |G|^{-1} \sum_{\chi \in \mathrm{Irr}(G)} |a_\chi|^2.$$

Proof. We have

$$\sum_{g \in G} P(g)^2 = \sum_{g \in G} P(g)\overline{P(g)} = |G|^{-2} \sum_{g \in G} \sum_{\chi} a_\chi \chi(g) \sum_{\psi} \overline{a_\psi}\overline{\psi(g)}$$

$$= |G|^{-2} \sum_{\chi,\psi} a_\chi \overline{a_\psi} \sum_{g \in G} \chi(g)\overline{\psi(g)}.$$

Using the orthogonality relations this gives

$$\sum_{g \in G} P(g)^2 = |G|^{-1} \sum_{\chi,\psi} a_\chi \overline{a_\psi} \delta_{\chi\psi} = |G|^{-1} \sum_{\chi} |a_\chi|^2.$$

■

Lemma 2.2 *We have*

$$\sum_{g\in G}(P(g)-|G|^{-1})^2 = |G|^{-1}\sum_{\chi\neq 1}|a_\chi|^2.$$

Proof. Using the previous lemma we have

$$\sum_{g\in G}(P(g)-|G|^{-1})^2 = \sum_{g\in G}P(g)^2 - 2|G|^{-1}\sum_{g\in G}P(g)+|G|^{-1} = |G|^{-1}(\sum_{\chi}|a_\chi|^2-1).$$

The result follows since $a_1 = 1$. ∎

Lemma 2.3 *We have*

$$||P-U||_1 \leq (\sum_{\chi\neq 1}|a_\chi|^2)^{1/2}.$$

Proof. By Cauchy-Schwarz inequality we have

$$(||P-U||_1)^2 = (\sum_{g\in G}|P(g)-|G|^{-1}|)^2 \leq |G|\sum_{g\in G}(P(g)-|G|^{-1})^2.$$

The result follows using the lemma above. ∎

Proof of Proposition 1.1

Let P be the commutator probability distribution on G. Then by (1) we have

$$a_\chi = \chi(1)^{-1} \text{ for all } \chi.$$

Applying the above lemma we obtain

$$||P-U||_1 \leq (\sum_{\chi\neq 1}\chi(1)^{-2})^{1/2} = (\zeta^G(2)-1)^{1/2}.$$

This completes the proof. ∎

Using the methods of this section we can easily show the following.

Proposition 2.4 *Let G be a finite group, and let $P_{x^2y^2}$ be the probability distribution associated with the word map from $G\times G$ to G induced by x^2y^2. Let R denote the set of real characters of G. Then*

$$||P_{x^2y^2}-U||_1 \leq (\sum_{\chi\in R}\chi(1)^{-2}-1)^{1/2} \leq (\zeta^G(2)-1)^{1/2}.$$

Proof. Let $\beta = \alpha_{x^2y^2} : G \times G \to G$. We use the well known formula (see e.g. [LiSh3], 3.1)

$$|\beta^{-1}(g)| = |G| \sum_{\chi \in R} \frac{\chi(g)}{\chi(1)}.$$

This means that

$$P_\beta = |G|^{-1} \sum_{\chi \in R} \chi(1)^{-1} \chi.$$

The result now follows from Lemma 2.3. ∎

3 Applications of the L_1 bound

The purpose of this section is to deduce results 1.2-1.4 from Proposition 1.1.

Proof of Corollary 1.2

Set $\delta = \delta(G)$. Then $||P_f - U||_1 \leq \delta$ by Proposition 1.1. Let C be the set of commutators in G, and let $D = G \setminus C$. Then

$$\delta \geq \sum_{g \in D} \left| P(g) - \frac{1}{|G|} \right| = |D|/|G|.$$

Therefore $|D| \leq \delta |G|$, so $|C| \geq (1 - \delta)|G|$, as required. ∎

The notion of almost equidistribution is naturally defined in the general setting of arbitrary functions between finite sets. Let X, Y be finite sets, and let $\epsilon > 0$. We say that a function $f : X \to Y$ is *ϵ-equidistributed* if there exists a subset $Y' \subseteq Y$ with the following properties:

(i) $|Y'| \geq |Y|(1 - \epsilon)$;

(ii) $\frac{|X|}{|Y|}(1 - \epsilon) \leq |f^{-1}(y)| \leq \frac{|X|}{|Y|}(1 + \epsilon)$ uniformly for all $y \in Y'$.

Recall that P_f is the probability distribution on Y induced by f. Let U be the uniform distribution on Y. We show that the L_1-distance

$$||P_f - U||_1 = \sum_{y \in Y} \left| P_f(y) - \frac{1}{|Y|} \right|$$

is small if and only if f is almost equidistributed. Indeed we have the following easy Lemma.

Lemma 3.1 *With the above notation we have*

(i) If f is ϵ-equidistributed then $||P_f - U||_1 \leq 4\epsilon$.

(ii) If $||P_f - U||_1 \leq \delta$, then f is $\sqrt{\delta}$-equidistributed.

Proof. (i) Assume that $f : X \to Y$ is ϵ-equidistributed, and let Y' be as in the definition above. Write $Y = Y' \cup Y''$ as a disjoint union. Then for any $y \in Y'$, $\left|P_f(y) - \frac{1}{|Y|}\right| \leq \frac{\epsilon}{|Y|}$, and thus

$$\left|\sum_{y\in Y'} P_f(y) - \sum_{y\in Y'} \frac{1}{|Y|}\right| \leq \sum_{y\in Y'} \left|P_f(y) - \frac{1}{|Y|}\right| \leq |Y'|\frac{\epsilon}{|Y|} \leq \epsilon.$$

However, $\sum_{y\in Y'} \frac{1}{|Y|} = \frac{|Y'|}{|Y|} \geq 1 - \epsilon$, therefore $\sum_{y\in Y'} P_f(y) \geq 1 - 2\epsilon$, so we deduce that $\sum_{y\in Y''} P_f(y) \leq 2\epsilon$. Therefore

$$\begin{aligned} ||P_f - U||_1 &= \sum_{y\in Y} \left|P_f(y) - \frac{1}{|Y|}\right| = \sum_{y\in Y'} \left|P_f(y) - \frac{1}{|Y|}\right| + \sum_{y\in Y''} \left|P_f(y) - \frac{1}{|Y|}\right| \\ &\leq \sum_{y\in Y'} \frac{\epsilon}{|Y|} + \sum_{y\in Y''} |P_f(y)| + \sum_{y\in Y''} \frac{1}{|Y|} \\ &\leq |Y'|\frac{\epsilon}{|Y|} + 2\epsilon + |Y''|\frac{1}{|Y|} \leq \epsilon + 2\epsilon + \epsilon = 4\epsilon. \end{aligned}$$

(ii) Assume that $||P_f - U||_1 \leq \delta$. Define $Y'' = \left\{y \in Y : \left|P_f(y) - \frac{1}{|Y|}\right| > \frac{\sqrt{\delta}}{|Y|}\right\}$. Then

$$\delta \geq \sum_{y\in Y} \left|P_f(y) - \frac{1}{|Y|}\right| \geq \sum_{y\in Y''} \left|P_f(y) - \frac{1}{|Y|}\right| > |Y''|\frac{\sqrt{\delta}}{|Y|}.$$

Therefore, $|Y''| < \sqrt{\delta}|Y|$. Take $Y' = Y \setminus Y''$. Then $|Y'| \geq |Y|(1 - \sqrt{\delta})$ and any $y \in Y'$ satisfies $\left|P_f(y) - \frac{1}{|Y|}\right| \leq \frac{\sqrt{\delta}}{|Y|}$. Thus, f is $\sqrt{\delta}$-equidistributed. ■

Proof of Theorem 1.3

This follows by combining Proposition 1.1 and part (ii) of Lemma 3.1.

■

The next result concerns measure preservation.

Proposition 3.2 *Let $f : X \to Y$ be ϵ-equidistributed.*

(i) If $Y_0 \subseteq Y$ then

$$\left|\frac{f^{-1}(Y_0)}{|X|} - \frac{|Y_0|}{|Y|}\right| \leq 3\epsilon.$$

(ii) If $X_0 \subseteq X$ then

$$\frac{|f(X_0)|}{|X|} \geq \frac{|X_0|}{|X|} - 3\epsilon.$$

Proof. Assume that $f : X \to Y$ is ϵ-equidistributed and let Y' be as in the definition above. Let $X' = f^{-1}(Y') \subseteq X$ be the inverse image of Y'. Then by part (i) and the lower bound in (ii) of the definition,

$$|X'| = \sum_{y \in Y'} |f^{-1}(y)| \geq |Y'| \frac{|X|}{|Y|}(1-\epsilon) \geq |X|(1-\epsilon)^2 \geq |X|(1-2\epsilon).$$

We conclude that

$$|f^{-1}(Y \setminus Y')| = |X| - |X'| \leq 2\epsilon |X|.$$

Now let $Y_0 \subseteq Y$. Then

$$|f^{-1}(Y_0)| \leq |f^{-1}(Y_0 \cap Y')| + |f^{-1}(Y \setminus Y')| \leq \sum_{y \in Y_0 \cap Y'} |f^{-1}(y)| + 2\epsilon |X|.$$

Using the upper bound in part (ii) of the definition, we see that

$$|f^{-1}(Y_0)| \leq |Y_0| \frac{|X|}{|Y|}(1+\epsilon) + 2\epsilon |X|.$$

Therefore

$$\frac{|f^{-1}(Y_0)|}{|X|} \leq \frac{|Y_0|}{|Y|}(1+\epsilon) + 2\epsilon \leq \frac{|Y_0|}{|Y|} + 3\epsilon.$$

On the other hand we have

$$\begin{aligned} |f^{-1}(Y_0)| \geq |f^{-1}(Y_0 \cap Y')| &= \sum_{y \in Y_0 \cap Y'} |f^{-1}(y)| \geq |Y \cap Y_0| \frac{|X|}{|Y|}(1-\epsilon) \\ &\geq (|Y_0| - \epsilon |Y|) \frac{|X|}{|Y|}(1-\epsilon). \end{aligned}$$

This yields

$$\frac{|f^{-1}(Y_0)|}{|X|} \geq \left(\frac{|Y_0|}{|Y|} - \epsilon \right)(1-\epsilon) \geq \frac{|Y_0|}{|Y|} - 2\epsilon.$$

This completes the proof of part (i) of the proposition.

Part (ii) now follows easily. Indeed, given $X_0 \subseteq X$, define $Y_0 = f(X_0)$. As above, we have

$$\frac{|f^{-1}(Y_0)|}{|X|} \leq \frac{|Y_0|}{|Y|} + 3\epsilon = \frac{|f(X_0)|}{|Y|} + 3\epsilon.$$

Since $X_0 \subseteq f^{-1}(Y_0)$ we obtain

$$\frac{|f(X_0)|}{|Y|} \geq \frac{|f^{-1}(Y_0)|}{|X|} - 3\epsilon \geq \frac{|X_0|}{|X|} - 3\epsilon.$$

■

Proof of Corollary 1.4

By Theorem 1.3 the commutator map $\alpha : G \times G \to G$ is $\epsilon(G)$-equidistributed. The corollary now follows immediately from Proposition 3.2.

■

4 Simple groups: character theoretic tools

In this section we introduce the main concepts and tools which are needed for our two proofs of Theorem 1.5.

The first result summarizes some of the properties of the Witten zeta function ζ^G defined in the Introduction. Our first proof of Theorem 1.5 follows by combining part (i) below with Theorem 1.3.

Theorem 4.1 *Let G be a finite simple group.*

(i) If $s > 1$ then $\zeta^G(s) \to 1$ as $|G| \to \infty$.

(ii) If $s > 2/3$ and $G \neq L_2(q)$ then $\zeta^G(s) \to 1$ as $|G| \to \infty$.

(iii) If $s > 0$ and $G = A_n$ then $\zeta^G(s) \to 1$ as $|G| \to \infty$. Moreover, $\zeta^G(s) = 1 + O(n^{-s})$.

(iv) If $G = G_r(q)$ then $\zeta^G(2) = 1 + O(q^{-r})$.

Indeed, parts (i) and (ii) are Theorem 1.1 of [LiSh4]. Part (iii) is Corollary 2.7 of [LiSh3].

To prove Part (iv) we note that

$$\zeta^G(2) - 1 \leq k(G)h(G)^{-2},$$

where $h(G)$ is the minimal degree of a non-trivial character of G. Suppose $G = G_r(q)$. Then

$$h(G) \geq c_1 q^r,$$

by Landazuri-Seitz [LaSe], and

$$k(G) \leq c_2 q^r$$

by Fulman-Guralnick [FG], where $c_1, c_2 > 0$ are absolute constants. Part (iv) now follows from the above inequalities.

The next result we need deals with regular semisimple elements in groups of Lie type.

Theorem 4.2 *Let $G = G_r(q)$ and let S be the set of regular semisimple elements in G.*

(i) There is an absolute constant a_1 such that

$$|S| \geq |G|(1 - a_1 q^{-1});$$

(ii) There is a number $f(r)$ depending only on r such that, if $g \in S$, then

$$|\chi(g)| \leq f(r)$$

for all $\chi \in \mathrm{Irr}(G)$.

Indeed, part (i) is a result of Guralnick and Lübeck [GL], while part (ii) is Lemma 4.4 of [Sh].

The next result we use deals with elements $g \in G_r(q)$ whose centralizer is not very large.

Theorem 4.3 *Let $G = G_r(q)$, and fix ϵ satisfying $0 < \epsilon < 2$. Let*

$$S(\epsilon) = \{g \in G : |C_G(g)| \leq q^{(3-\epsilon)r}\}.$$

Then we have

(i) $|S(\epsilon)| \geq |G|(1 - a_2 q^{-(2-\epsilon)r})$, where a_2 is an absolute constant;

(ii) There is a number $r_1(\epsilon)$ such that, if $r \geq r_1(\epsilon)$, and $g \in S(\epsilon)$, then

$$\sum_{1 \neq \chi \in \mathrm{Irr}(G)} \frac{|\chi(g)|}{\chi(1)} \to 0 \text{ as } |G| \to \infty.$$

Indeed part (i) follows from Corollary 5.4 of [Sh], while part (ii) is Proposition 4.7 of [Sh].

We conclude this section quoting a useful result of Müller and Schlage-Puchta on character values for symmetric groups. See part (i) of Theorem A in [MS].

Theorem 4.4 *Let $g \in S_n$ be a permutation with f fixed points. Define*

$$\delta = ((1 - 1/\log n)^{-1} \frac{12 \log n}{\log (n/f)} + 18)^{-1}.$$

Then we have

$$|\chi(g)| \leq \chi(1)^{1-\delta}$$

for all $\chi \in \mathrm{Irr}(S_n)$.

Remark

By Theorem 4.1 we have $\zeta^G(2) - 1 = O(n^{-2})$ where $G = A_n$, and $\zeta^G(2) - 1 = O(q^{-r})$ where $G = G_r(q)$. This shows that

$$\epsilon(A_n) = O(n^{-1/2}) \text{ and } \epsilon(G_r(q)) = O(q^{-r/4}).$$

5 Theorem 1.5: constructive proof

In this section we prove Theorem 1.5 in a constructive manner, providing the subsets $S \subset G$ with the required properties. We need some notation.

For $g \in G$ let

$$\Delta(g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G)} \frac{\chi(g)}{\chi(1)},$$

and

$$E(g) = \sum_{1 \neq \chi \in \mathrm{Irr}(G)} \frac{|\chi(g)|}{\chi(1)}.$$

Then $N(g) = |G|(1 + \Delta(g))$ and $|\Delta(g)| \leq E(g)$.

To prove Theorem 1.5 constructively it therefore suffices to find subsets $S \subseteq G$ consisting of almost all elements of G such that $E(g) = o(1)$ for all $g \in S$. We will see below that this strategy works out for some (but not all) finite simple groups G. For the remaining groups we will show that $|\Delta(g)| = o(1)$, which also suffices.

Our study of $E(g)$ and $\Delta(g)$ is based on recent detailed results on character degrees and values quoted in Section 4.

Case 1. $G = L_2(q)$.

The character table of $L_2(q)$ is well known (see for instance [Do]), so $\Delta(g)$ and $N(g)$ can be computed for all $g \in G$ using (1). We summarize the result below.

Proposition 5.1 *Let $q = p^n$ be a prime power, then in the group $L_2(q)$,*

1. *If a is an element of order $\frac{q-1}{2}$ (when q is odd) or order $q-1$ (when q is even) then $\Delta(a^l) = \frac{1}{q} + \frac{\alpha(q,l)}{q+1}$ where*

$$\alpha(q,l) = \begin{cases} 2 & \text{if } q \equiv 1 \pmod 4, \quad l \text{ even} \\ -4 & \text{if } q \equiv 1 \pmod 4, \quad l \text{ odd} \\ -1 & \text{if } q \equiv 3 \pmod 4 \text{ or } q \text{ is even} \end{cases}$$

2. *If b is an element of order $\frac{q+1}{2}$ (when q is odd) or order $q+1$ (when q is even) then $\Delta(b^m) = -\frac{1}{q} + \frac{\beta(q,m)}{q-1}$ where*

$$\beta(q,m) = \begin{cases} 1 & \text{if } q \equiv 1 \pmod 4 \text{ or } q \text{ is even} \\ -2 & \text{if } q \equiv 3 \pmod 4, \quad m \text{ even} \\ 4 & \text{if } q \equiv 3 \pmod 4, \quad m \text{ odd} \end{cases}$$

3. *If c is an element of order p then*

$$\Delta(c) = \begin{cases} \frac{1}{2(q+1)} & \text{if } q \equiv 1 \pmod 4 \\ -\frac{1}{q+1} - \frac{3}{2(q-1)} & \text{if } q \equiv 3 \pmod 4 \\ -\frac{3}{2(q+1)} - \frac{1}{2(q-1)} & \text{if } q \text{ is even} \end{cases}$$

Letting S be all non-identity elements of G we find by Proposition 5.1 that for $g \in S$,

$$|\Delta(g)| \leq 5/q = o(1).$$

This implies the required conclusion.

Let $\epsilon, r_1(\epsilon)$ be as in Theorem 4.3, and set $b = r_1(1)$.

Case 2. $G = G_r(q)$, where $r < b$, and $G \neq L_2(q)$.

Let S be the set of regular semisimple elements of G. Since the rank r is bounded, we have $q \to \infty$ as $|G| \to \infty$.

By Theorem 4.2(i) above we have

$$|S| \geq |G|(1 - a_1 q^{-1}) = |G|(1 - o(1)).$$

By part (ii) of Theorem 4.2 there is a number $f(r)$ depending only on r such that

$$|\chi(g)| \leq f(r)$$

for all $\chi \in \operatorname{Irr} G$ and $g \in S$. This yields

$$E(g) \leq \sum_{\chi \neq 1} f(r)\chi(1)^{-1} = f(r)(\zeta^G(1) - 1).$$

Since $G \neq L_2(q)$ we have $\zeta^G(1) \to 1$ as $|G| \to \infty$ (see part (ii) of Theorem 4.1). This yields $E(g) = o(1)$ uniformly for all $g \in S$, proving the result in this case.

Case 3. $G = G_r(q)$ where $r \geq b$.

We adopt the notation of Theorem 4.3 and apply this result with $\epsilon = 1$. Part (i) then yields

$$|S(1)| \geq |G|(1 - a_2 q^{-r}).$$

Now let $g \in S(1)$. Since $r \geq b = r_1(1)$, part (ii) of Theorem 4.3 shows that $E(g) = o(1)$, and again the conclusion follows with $S = S(1)$.

Case 4. $G = A_n$.

Let S be the set of permutations in A_n with at most $\sqrt{n}$ fixed points.

It is easy to see that the probability that a permutation $g \in A_n$ has at least f fixed points is at most $2/f!$. This implies that

$$|S| \geq |A_n|(1 - 2/[\sqrt{n}]!) = |A_n|(1 - o(1)).$$

Now set

$$\delta = \frac{1}{43},$$

and let $g \in S$. Using Theorem 4.4 above we see that, for n large, we have

$$|\chi(g)| \leq \chi(1)^{1-\delta} \tag{2}$$

for all $\chi \in \mathrm{Irr}(S_n)$.

For each irreducible character χ of S_n, either $\chi \downarrow A_n$ is irreducible, or $\chi \downarrow A_n = \chi_1 + \chi_2$, a sum of two irreducible characters of degree $\chi(1)/2$. All irreducible characters of A_n occur in this way.

In the latter case, note that

$$\frac{\chi_1(g)}{\chi_1(1)} + \frac{\chi_2(g)}{\chi_2(1)} = 2\frac{\chi(g)}{\chi(1)}.$$

This implies that

$$\left| \sum_{1 \neq \chi \in \mathrm{Irr}(A_n)} \frac{\chi(g)}{\chi(1)} \right| \leq 2 \sum_{\chi \in \mathrm{Irr}(S_n), \chi(1) > 1} \frac{|\chi(g)|}{\chi(1)} \leq 2 \sum_{\chi \in \mathrm{Irr}(S_n), \chi(1) > 1} \chi(1)^{-\delta},$$

where the last inequality follows from (2). We conclude that, in A_n we have

$$|\Delta(g)| \leq 2(\zeta^{S_n}(\delta) - 2).$$

By Theorem 1.1 in [LiSh3], $\zeta^{S_n}(\delta) = 2 + O(n^{-\delta})$. This yields

$$|\Delta(g)| = O(n^{-1/43}) \to 0 \text{ as } n \to \infty.$$

This completes the case $G = A_n$ and the constructive proof of Theorem 1.5. ∎

Remark

The proof above yields explicit lower bounds on $|S|$. In some cases they may be further improved. Indeed, fix any $\epsilon > 0$. Then for $G = G_r(q)$ and $r \geq r_1(\epsilon)$ we can use Theorem 4.3 to construct S satisfying

$$|S| \geq |G|(1 - a_2 q^{-(2-\epsilon)r}).$$

For $G = A_n$ we may take S as all even permutations with at most $n^{1-\epsilon}$ fixed points, obtaining

$$|S| \geq |A_n|(1 - 2/[n^{1-\epsilon}]!) \geq |A_n|(1 - n^{-n^{\gamma}}),$$

where $\gamma < 1$ is arbitrarily close to 1.

6 Applications to T-systems

In this section we focus on the various applications and prove results 1.7-1.9.

Proof of Theorem 1.7

This is an immediate consequence of Corollary 1.6(iii), in view of the fact that almost all pairs of finite simple groups are generating pairs (see [LiSh1] and the references therein).

■

We now turn to the main applications, involving T_2-systems.

Proof of Theorem 1.8

It suffices to prove the second (quantitative) assertion in the theorem.

Higman's Lemma states that for $k = d(G) = 2$, the union of the conjugacy classes under $\mathrm{Aut}(G)$ of the commutators $[g_1, g_2]$ and $[g_2, g_1]$ is an invariant of the T_2-system of (g_1, g_2) (see [GP, Ne, P]). This lemma becomes useful when dealing with T_2-systems of finite simple groups G.

By Theorem 1.7 there is a subset $S \subset G$ such that $S = |G|(1-o(1))$, and every $g \in S$ can be written as $g = [g_1, g_2]$ where $(g_1, g_2) \in V_2(G)$ (namely g_1, g_2 generate G, thus giving rise to a T_2-system).

Let $k(S)$ denote the number of distinct unions $C \cup C^{-1}$ where C is an $\mathrm{Aut}(G)$-conjugacy class of an element of S. Then the number of T_2-systems in G is at least $k(S)$.

Suppose first $G = A_n$. Then $\mathrm{Aut}(G) = S_n$ and $C = C^{-1}$ for any S_n-class C. Given $\epsilon > 0$ fix $\delta > 0$ such that $\delta < \epsilon$. We may assume n is sufficiently large (given ϵ). By a result of Erdős and Turán (see [ET]), the order of almost all permutations in S_n is at least $n^{(1/2-\delta)\log n}$. Intersecting S with the set of permutations with this order restriction we may therefore assume that each $g \in S$ has order at least $n^{(1/2-\delta)\log n}$, while we still have $|S| = |A_n|(1 - o(1)) \geq n!/3$.

Now, if $g \in S$, then we have

$$|C_{S_n}(g)| \geq |g| \geq n^{(1/2-\delta)\log n}.$$

Thus

$$|g^{S_n}| \leq n! \cdot n^{-(1/2-\delta)\log n}.$$

Since the union of the S_n-classes of $g \in S$ covers S we have

$$k(S) \cdot n! \cdot n^{-(1/2-\delta)\log n} \geq |S| \geq n!/3.$$

This yields

$$k(S) \geq n^{(1/2-\delta)\log n}/3 \geq n^{(1/2-\epsilon)\log n},$$

proving the result for A_n.

Now suppose $G = G_r(q)$.

Define

$$c(G) = \min\{|C_G(g)| : g \in G\}.$$

Let $\mathrm{Out}(G) = \mathrm{Aut}(G)/G$ and let $g \in G$. Then

$$|g^{\mathrm{Aut}(G)}| = |\mathrm{Aut}(G)|/|C_{\mathrm{Aut}(G)}(g)| \leq |\mathrm{Aut}(G)|/c(G).$$

In particular the size of each union $C \cup C^{-1}$ where $C = g^{\mathrm{Aut}(G)}$ and $g \in S$ is at most $2|\mathrm{Aut}(G)|/c(G)$. Since there are $k(S)$ such unions, and their union covers S, we see that

$$k(S) \cdot 2|\mathrm{Aut}(G)|/c(G) \geq |S| \geq |G|(1 - o(1)).$$

This yields

$$k(S) \geq (1/2 - o(1))c(G)/|\mathrm{Out}(G)|.$$

By results of Fulman and Guralnick [FG] there is an absolute constant $c_1 > 0$ such that

$$c(G_r(q)) \geq c_1 q^r / \log q.$$

The structure of $\mathrm{Out}(G)$ is known, and assuming $q = p^f$ (p prime) we have

$$|\mathrm{Out}(G)| \leq c_2 r f \leq c_3 r \log q$$

for some (small) constant c_3. The above inequalities yield

$$k(S) \geq c_4 q^r r^{-1} (\log q)^{-2}.$$

This completes the proof of Theorem 1.8. ∎

Remarks

1. Note that, by [ET], almost all permutations in S_n have at most $(1 + \delta)\log n$ cycles. These permutations split into at most $n^{(1+\epsilon)\log n}$ conjugacy classes (since this number bounds the number of partitions of n into at most $(1 + \delta)\log n$ parts, where $\epsilon > \delta$). Thus, although $k(S_n) = p(n) > c^{\sqrt{n}}$, a

union of just $n^{(1+\epsilon)\log n}$ conjugacy classes covers almost all of S_n. This shows that our lower bound on $k(S)$ in the proof above is essentially best possible.

2. If $G = G_r(q)$ the proof above shows that the number of T_2-systems in G is at least $(1/2-o(1))c(G)/|\operatorname{Out}(G)|$. This produces specific bounds which are slightly better than the general one stated in the theorem. For example, it follows that the number of T_2-systems in $L_2(p)$ is at least $(1/8 - o(1))p$.

We conclude this section by briefly discussing the PRA graph $\Gamma_2(G)$. An elementary calculation shows that the conjugacy class in G of the commutator $[g_1, g_2]$ is an invariant of the connected component of (g_1, g_2) in $\Gamma_2(G)$. Arguments similar to the proof of Theorem 1.8 show that the number of components of this graph is at least $(1 - o(1))c(G)$. This shows that $\Gamma_2(G_r(q))$ has at least $aq^r/\log q$ connected components, giving a better lower bound in Corollary 1.9. The lower bound for A_n remains the same.

7 Equidistribution revisited

Any word $w(x_1, \ldots, x_m) \in F_m$ gives rise to a word map $\alpha_w : G^m \to G$. Word maps on algebraic groups and on finite simple groups have been the subject of active investigations in recent years, see [Bo], [LiSh2], [La], [Sh] and [LaSh].

It is interesting to find out which words w have the remarkable property of the commutator map, namely that α_w is almost equidistributed (namely $o(1)$-equidistributed) on all finite simple groups G.

In general this is highly unexpected. For example, power words $w = x_1^k$ ($k \geq 2$), or words which are proper powers $w = w_1^k$, have image much smaller than $(1 - o(1))|G|$ for infinite families of finite simple groups G, hence their associated maps are not almost equidistributed.

However, we do obtain positive result for some more words.

Theorem 7.1 *Let G be a finite simple group, and let $\beta : G \times G \to G$ be the map given by $\beta(x, y) = x^2y^2$. Then there is a subset $S \subset G$ with $|S| = (1 - o(1))|G|$ such that $|\beta^{-1}(g)| = (1 + o(1))|G|$ for all $g \in S$.*

Proof. By Proposition 2.4 $||P_\beta - U||_1 \leq \delta(G)$. Lemma 3.1(ii) now shows that β is $\epsilon(G)$-distributed. Finally, by Theorem 4.1(i), $\epsilon(G) = o(1)$. The result follows. ∎

To obtain more positive results we need some preparations.

The following two lemmas show that the property of almost equidistribution behaves well under direct products and compositions. The proofs use the L_1 notation, which is more natural.

Lemma 7.2 *Let X_1, X_2, Y_1, Y_2 be finite sets and let $\delta_1, \delta_2 > 0$. For $i = 1, 2$ denote by U_i the uniform distribution on Y_i and assume that $f_i : X_i \to Y_i$ satisfies $||P_{f_i} - U_i||_1 \leq \delta_i$.*

Then the function $f = f_1 \times f_2 : X_1 \times X_2 \to Y_1 \times Y_2$, which is defined by

$$f(x_1, x_2) = (f_1(x_1), f_2(x_2)) \text{ for } x_1 \in X_1, x_2 \in X_2,$$

satisfies $||P_f - U||_1 \leq \delta_1 + \delta_2$, where U is the uniform distribution on $Y_1 \times Y_2$.

Proof. We have $P_{f_1 \times f_2}(y_1, y_2) = P_{f_1}(y_1) P_{f_2}(y_2)$ for any $(y_1, y_2) \in Y_1 \times Y_2$. Thus,

$$\begin{aligned}
||P_{f_1 \times f_2} - U||_1 &= \sum_{(y_1,y_2) \in Y_1 \times Y_2} \left| P_{f_1}(y_1) P_{f_2}(y_2) - \frac{1}{|Y_1|} \frac{1}{|Y_2|} \right| \\
&\leq \sum_{y_2 \in Y_2} P_{f_2}(y_2) \sum_{y_1 \in Y_1} \left| P_{f_1}(y_1) - \frac{1}{|Y_1|} \right| \\
&\quad + \sum_{y_1 \in Y_1} \frac{1}{|Y_1|} \sum_{y_2 \in Y_2} \left| P_{f_2}(y_2) - \frac{1}{|Y_2|} \right| \\
&\leq \sum_{y_2 \in Y_2} P_{f_2}(y_2) \delta_1 + \sum_{y_1 \in Y_1} \frac{1}{|Y_1|} \delta_2 = \delta_1 + \delta_2.
\end{aligned}$$

■

Lemma 7.3 *Let X, Y, Z be finite sets, and let $\delta_1, \delta_2 > 0$. Denote by U_Y and U_Z the uniform distributions on Y and Z respectively. Assume that $f_1 : X \to Y$ and $f_2 : Y \to Z$ satisfy $||P_{f_1} - U_Y||_1 \leq \delta_1$ and $||P_{f_2} - U_Z||_1 \leq \delta_2$. Then their composition $f = f_2 \circ f_1 : X \to Z$ satisfies $||P_f - U_Z||_1 \leq \delta_1 + \delta_2$.*

Proof. Note that for any $z \in Z$,

$$P_{f_2 \circ f_1}(z) = \frac{|f_1^{-1}(f_2^{-1}(z))|}{|X|} = \sum_{f_2(y)=z} \frac{|f_1^{-1}(y)|}{|X|} = \sum_{f_2(y)=z} P_{f_1}(y).$$

Thus, by the assumption on f_1,

$$\begin{aligned}
\sum_{z \in Z} \left| P_{f_2 \circ f_1}(z) - \sum_{f_2(y)=z} \frac{1}{|Y|} \right| &\leq \sum_{z \in Z} \sum_{f_2(y)=z} \left| P_{f_1}(y) - \frac{1}{|Y|} \right| \\
&= \sum_{y \in Y} \left| P_{f_1}(y) - \frac{1}{|Y|} \right| \leq \delta_1,
\end{aligned}$$

and by the assumption on f_2,

$$\sum_{z\in Z}\left|\sum_{f_2(y)=z}\frac{1}{|Y|}-\frac{1}{|Z|}\right| = \sum_{z\in Z}\left|\frac{|f_2^{-1}(z)|}{|Y|}-\frac{1}{|Z|}\right| = \sum_{z\in Z}\left|P_{f_2}(z)-\frac{1}{|Z|}\right| \leq \delta_2.$$

Therefore,

$$\begin{aligned}\|P_{f_2\circ f_1}-U_Z\|_1 &= \sum_{z\in Z}\left|P_{f_2\circ f_1}(z)-\frac{1}{|Z|}\right| \\ &\leq \sum_{z\in Z}\left(\left|P_{f_2\circ f_1}(z)-\sum_{f_2(y)=z}\frac{1}{|Y|}\right|+\left|\sum_{f_2(y)=z}\frac{1}{|Y|}-\frac{1}{|Z|}\right|\right) \\ &\leq \delta_1+\delta_2.\end{aligned}$$

■

Let $\delta > 0$ and let $f_1, \ldots, f_n$ be functions between finite sets, such that each one of them satisfies $||f_i - U_i||_1 \leq \delta$, where U_i is the uniform distribution on the range of f_i. Assume that we apply a finite number of steps m, such that each step is a composition or a direct product of two functions, and obtain a new function f. From the lemmas above we readily deduce that f satisfies $||f - U||_1 \leq m\delta$, where U is the uniform distribution on the range of f. By Lemma 3.1(ii), f is therefore $\sqrt{m\delta}$-equidistributed.

Using the above observation and Proposition 1.1 we now easily obtain the following.

Theorem 7.4 *Let G be a finite group. Let $m \geq 1$ and $w = [x_1, \ldots, x_m]$, an m-fold commutator in any arrangements of brackets. Then the associated word map $\alpha_w : G^m \to G$ is $\gamma(G)$-equidistributed, where $\gamma(G) = (m-1)^{1/2}(\zeta^G(2)-1)^{1/4}$.*

In particular, if G is simple, then α_w is almost equidistributed as $|G| \to \infty$.

Combining Theorems 7.1 and 7.4 with Proposition 3.2 we see that word maps α_w associated with $w = x^2y^2$, or with any m-fold commutator w, are almost measure preserving on finite simple groups.

In particular, using the fact that almost all pairs are generating pairs, we obtain

Corollary 7.5 *Almost all elements g of a finite simple group G can be obtained as $g = [g_1, \ldots, g_m]$, an m-fold commutator in any given arrangements of brackets, where $g_1, \ldots, g_m \in G$ satisfy $\langle g_i, g_j\rangle = G$ for all $i \neq j$.*

We can add various extra conditions on the m-tuple $(g_1, \ldots, g_m)$ above, provided they hold with probability tending to 1. For example, given any non-trivial words $w_1, \ldots, w_k \in F_m$ we can require that $w_i(g_1, \ldots, g_m) \neq 1$ for all $i = 1, \ldots, k$. Indeed, it is proved in [DPSSh] that, if G is a finite simple group and $1 \neq w \in F_m$, then, as $|G| \to \infty$, almost all m-tuples $(g_1, \ldots, g_m) \in G^m$ satisfy $w(g_1, \ldots, g_m) \neq 1$.

For instance, one now easily deduces that almost all elements g of a finite simple group G can be expressed as $g = [[g_1, g_2], [g_3, g_4]]$ where $\langle g_i, g_j \rangle = G$ for $i \neq j$, and the orders of $g_1, \ldots, g_4$ are as large as we want.

References

[B] L. Babai, The probability of generating the symmetric group, *J. Comb. Th. Ser. A* **52** (1989), 148–153.

[BP] L. Babai, I. Pak, Strong bias of group generators: an obstacle to the "product replacement algorithm", *Proc. Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms* (2000).

[Bo] A. Borel, On free subgroups of semisimple groups, *Enseign. Math.* **29** (1983), 151–164.

[CLMNO] F. Celler, C.R. Leedham-Green, S. Murray, A. Niemeyer, E.A. O'Brien, Generating random elements of a finite group, *Comm. Alg.* **23** (1995), 4931–4948.

[Di] J.D. Dixon, The probability of generating the symmetric group, *Math. Z.* **110** (1969), 199–205.

[DPSSh] J.D. Dixon, L. Pyber, Á. Seress, A. Shalev, Residual properties of free groups and probabilistic methods, *J. reine angew. Math. (Crelle's)* **556** (2003), 159–172.

[Do] L. Dornhoff, *Group Representation Theory, Part A*, Marcel Dekker, 1971.

[Du1] M.J. Dunwoody, On T-systems of groups, *J. Austral. Math. Soc.* **3**, (1963), 172–179.

[Du2] M.J. Dunwoody, Nielsen transformations, *in: Computation Problems in Abstract Algebra*, Pergamon, Oxford, 1970, 45–46.

[EG] E.W. Ellers and N. Gordeev, On conjectures of J. Thompson and O. Ore, *Trans. Amer. Math. Soc.* **350**, 3657–3671.

[ET] P. Erdős and P. Turán, On some problems of a statistical group theory. I, *Z. Wahrscheinlichkeitstheorie Verw. Gabiete* **4** (1965), 175–186.

[E] M.J. Evans, Ph.D. Thesis, University of Wales, 1985.

[E2] M.J. Evans, T-systems of certain finite simple groups, *Math. Proc. Cambridge Philos. Soc.* **113** (1993), 9–22.

[FG] J. Fulman and R.M. Guralnick, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups, Preprint.

[Ga] P.X. Gallagher, The generation of the lower central series, *Canad. J. Math.* **17** (1965), 405–410.

[GaP] A. Gamburd, I. Pak, Expansion of product replacement graphs, *Combinatorica* **26** (2006), no. 4, 411–429.

[Gi] R. Gilman, Finite quotients of the automorphism group of a free group, *Canad. J. Math.* **29** (1977), 541–551.

[Go] R. Gow, Commutators in finite simple groups of Lie type, *Bull. London Math. Soc.* **32** (2000), 311–315.

[GL] R.M. Guralnick and F. Lübeck, On p-singular elements in Chevalley groups in characteristic p, *Groups and computation, III* (Columbus, OH, 1999), 169–182, Ohio State Univ. Math. Res. Inst. Publ., 8, de Gruyter, Berlin, 2001.

[GP] R.M. Guralnick and I. Pak, On a question of B.H. Neumann, *Proc. Amer. Math. Soc.* **131** (2002), 2021–2025.

[Ja] A. Jaikin-Zapirain, Zeta function of representations of compact p-adic analytic groups, *J. Amer. Math. Soc.* **19** (2006), 91–118.

[KL] W.M. Kantor and A. Lubotzky, The probability of generating a finite classical group, *Geom. Ded.* **36** (1990), 67–87.

[LaSe] V. Landazuri and G.M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* **32** (1974), 418-443.

[La] M. Larsen, Word maps have large image, *Israel J. Math.* **139** (2004), 149–156.

[LaSh] M. Larsen and A. Shalev, Word maps and Waring type problems, Preprint, 2007.

[LiSh1] M.W. Liebeck and A. Shalev, The probability of generating a finite simple group, *Geom. Ded.* **56** (1995), 103–113.

[LiSh2] M.W. Liebeck and A. Shalev, Diameters of finite simple groups: sharp bounds and applications, *Annals of Math.* **154** (2001), 383–406.

[LiSh3] M.W. Liebeck and A. Shalev, Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks, *J. Algebra* **276** (2004), 552–601.

[LiSh4] M.W. Liebeck and A. Shalev, Fuchsian groups, finite simple groups, and representation varieties, *Invent. Math.* **159** (2005), 317–367.

[LiSh5] M.W. Liebeck and A. Shalev, Character degrees and random walks in finite groups of Lie type, *Proc. London Math. Soc.* **90** (2005), 61–86.

[LM] A. Lubotzky and B. Martin, Polynomial representation growth and the congruence subgroup problem, *Israel J. Math.* 144 (2004), 293–316.

[LP] A. Lubotzky, I. Pak, The product replacement algorithm and Kazhdan's property (T), *Journal of the AMS* **52** (2000), 5525–5561.

[MS] T.W. Müller and J-C. Schlage-Puchta, Character Theory of Symmetric Groups, Subgroup Growth of Fuchsian Groups, and Random Walks, preprint, arXiv: math.GR/0305260.

[Ne] B.H. Neumann, On a question of Gaschütz, *Arch. Math.* **7** (1956), 87–90.

[NN] B.H. Neumann, H. Neumann, Zwei klassen charakteristischer untergruppen und ihre faktorgruppen, *Math. Nachr.* **4** (1951), 106–125.

[O] O.O. Ore, Some remarks on commutators, *Proc. Amer. Soc.* **272** (1951), 307–314.

[P] I. Pak, What do we know about the product replacement algorithm?, in: *Groups and computation III*, eds: Kantor and Seress, de Gruyter, Berlin, 2000, pp. 301–347.

[Sh] A. Shalev, Word maps, conjugacy classes, and a non-commutative Waring-type theorem, to appear in *Annals of Math.*

[Wi] J.S. Wilson, On simple pseudofinite groups, *J. London Math. Soc.* **51** (1995), 471–490.

[Wit] E. Witten, On quantum gauge theories in two dimensions, *Comm. Math. Phys.* **141** (1991), 153–209.